



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 138997

TO: Cynthia Wilder
Location: REM/2A35/2C18
Wednesday, April 20, 2005
Art Unit: 1637

Case Serial Number: 09/529397

From: David Schreiber
Location: Biotech-Chem Library
Remsen E01A61
Phone: 571-272-2526

David.Schreiber@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 13:28:09 ; Search time 427 Seconds
(without alignments)
1386.356 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 999gagagagaaucaucg.....uagcagcagaagcucucg 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980bs.*
2: Geneseqn1990bs.*
3: Geneseqn2000bs.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	100	3	Aaz99048 RNA aptam
2	100	100.0	100	12	ADJ10063 Synthetic
3	100	100.0	100	12	ADJ10064 Synthetic
4	100	100.0	200	12	ADJ10065 Synthetic
5	90	90.0	90	3	Aaz99050 RNA aptam
6	84.4	84.4	98	12	ADJ10053 RNA aptam
7	80	80.0	80	3	Aaz99051 RNA aptam
8	50.6	50.6	147	12	ADJ10066 Synthetic
9	45	45.0	45	3	Aaz99076 Ras target
c 10	30.6	30.6	748	3	Aaz99076 Human ORF
c 11	30.6	30.6	748	6	ABN24508 Human ORF
12	30.6	30.6	15776	8	Aaz99083 Propionib
13	30.6	30.6	15776	8	ACF64512 Propionib
c 14	30.4	30.4	366	9	ADZ29321 DNA encod
15	30	30.0	98	3	Aaz99049 RNA aptam
16	30	30.0	90616	12	ADQ97596 Mouse can
17	29.8	29.8	4192	3	Aaz99049 RNA aptam
18	29.8	29.8	4285	4	AAE85391 Nucleotid
19	29.8	29.8	4285	6	ABK88484 Poplar pr
20	29.8	29.8	4285	9	ACA62517 Poplar ho

21	29	29.0	29	3	Aaz99084	Aaz99084 Ras target
c 22	29	29.0	29	12	ADJ10055	ADJ10055 Synthetic
c 23	28.8	28.8	675	6	ACA52761	ACA52761 Prokaryot
c 24	28.6	28.6	2000	6	ABZ15103	ABZ15103 Arabidops
c 25	28.2	28.2	1161	5	ABV24321	ABV24321 Human pro
c 26	28.2	28.2	1665	2	AAx58401	AAx58401 Vicia sat
c 27	28	28.0	29	12	ADJ10058	ADJ10058 PCR prime
c 28	28	28.0	29	12	ADJ10057	ADJ10057 PCR prime
c 29	28	28.0	29	12	ADJ10056	ADJ10056 PCR prime
c 30	28	28.0	110000	6	ABX08336	Continuation (2 of
c 31	28	28.0	110000	12	ADJ25985	Continuation (2 of
c 32	28	28.0	110000	12	ADN97989	Continuation (2 of
c 33	28	28.0	110000	12	ADO50281	Continuation (2 of
34	27.8	27.8	1093	13	ADS47230	ADs47230 Bacterial
c 35	27.8	27.8	99291	13	ABD33549	ABD33549 Human can
36	27.6	27.6	732	4	AAI94947	AAI94947 Human neu
37	27.6	27.6	732	8	ABT42823	ABt42823 Human neu
c 38	27.4	27.4	3791	2	AAT63575	Aat63575 Chicken b
c 39	27	27.0	27	3	Aaz99087	Aaz99087 Ras target
c 40	27	27.0	29	12	ADJ10059	ADJ10059 PCR prime
c 41	27	27.0	208	3	AAC17425	Aac17425 Human sec
42	27	27.0	522	3	AAA69206	AAA69206 Bacteriop
43	27	27.0	56506	3	AAA69168	AAA69168 Bacteriop
c 44	26.8	26.8	415	3	AAC21986	Aac21986 Human sec
45	26.8	26.8	1116	5	AAH51835	AAH51835 HIV prote

ALIGNMENTS

RESULT 1
Aaz99048
ID Aaz99048 standard; RNA; 100 BP.
XX
AC Aaz99048;
XX
DT 21-JUN-2000 (first entry)
XX
DE RNA aptamer #25 for binding Ras target protein.

XX Ras target protein; malignant tumour; signal transduction regulation;
XX cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX Homo sapiens.

XX HQ200009684-A1.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-JP004399.
XX
PR 14-AUG-1998; 98JP-00242596.
PR 24-NOV-1998; 98JP-00333284.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Yokoyama S, Hirao I, Sakamoto K;
XX WPI; 2000-224330/19.

(PD) Priority document

XX Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras target protein like Raf-1, useful in drug compositions to treat and diagnose malignant tumors and inflammation.

XX Claim 6; Page 41; 59pp; Japanese.

XX The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1. RNA aptamer (Aaz99048-299051) based on these sequences are useful in the treatment and diagnosis of malignant tumours and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumours and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal

11/24/98

CC causing proliferation or differentiation of cells
XX
SQ Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;
Query Match 100.0%; Score 100; DB 3; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.7e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGAGUGGAGGAUUAUCGAGGCAUAUGUCGACUCCGUCUCCUUAACAGUUUAA 60
Db 1 GGGAGUGGAGGAUUAUCGAGGCAUAUGUCGACUCCGUCUCCUUAACAGUUUAA 60
QY 61 AUUGGUUUUAGCAUAGCCUUAAGCGACAGCAAGCUUCUGC 100
Db 61 AUUGGUUUUAGCAUAGCCUUAAGCGACAGCAAGCUUCUGC 100
RESULT 2
ADJ10063
ID ADJ10063 standard; DNA; 100 BP.
XX
AC ADJ10063;
XX
DT 17-JUN-2004 (first entry)
XX
DE Synthetic RNA 9A (100-mer).
XX
KW ss; unnatural base; 5'-substituted-2-oxo(1H)-pyridin-3-yl;
KW photoactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS Synthetic.
XX
PN WO2004007713-A1.
XX
PD 22-JAN-2004.
XX
PF 28-FEB-2003; 2003WO-JP002342.
XX
PR 17-JUL-2002; 2002JP-00208569.
XX
PA (RIKE) RIKEN KK.
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Hirao I, Yokoyama S, Hirao M, Mitsui T;
XX
DR WPI; 2004-122944/12.
XX
PT Nucleosides or nucleotides comprising novel base particularly 5-substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.
XX
PS Disclosure; Fig 10; 78pp; Japanese.
XX
CC This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)-pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photoactive group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as

CC causing proliferation or differentiation of cells
XX
SQ Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;
Query Match 100.0%; Score 100; DB 12; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.7e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGAGUGGAGGAUUAUCGAGGCAUAUGUCGACUCCGUCUCCUUAACAGUUUAA 60
Db 1 GGGAGUGGAGGAUUAUCGAGGCAUAUGUCGACUCCGUCUCCUUAACAGUUUAA 60
QY 61 AUUGGUUUUAGCAUAGCCUUAAGCGACAGCAAGCUUCUGC 100
Db 61 AUUGGUUUUAGCAUAGCCUUAAGCGACAGCAAGCUUCUGC 100
RESULT 3
ADJ10064
ID ADJ10064 standard; DNA; 100 BP.
XX
AC ADJ10064;
XX
DT 17-JUN-2004 (first entry)
XX
DE Synthetic RNA 9A(51y87) aptamer (100-mer).
XX
KW ss; unnatural base; 5'-substituted-2-oxo(1H)-pyridin-3-yl;
KW photoactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS Synthetic.
XX
FH Key modified_base 87
FT /tag= a
FT /mod_base= OTHER
FT /note= "OTHER= 51y"
XX
PN WO2004007713-A1.
XX
PD 22-JAN-2004.
XX
PF 28-FEB-2003; 2003WO-JP002342.
XX
PR 17-JUL-2002; 2002JP-00208569.
XX
PA (RIKE) RIKEN KK.
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Hirao I, Yokoyama S, Hirao M, Mitsui T;
XX
DR WPI; 2004-122944/12.
XX
PT Nucleosides or nucleotides comprising novel base particularly 5-substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.
XX
PS Disclosure; Fig 10; 78pp; Japanese.
XX
CC This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)-pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photoactive group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as

Query Match 90.0%; Score 90; DB 3; Length 90;
Best Local Similarity 100.0%; Pred. No. 9e-22; Mismatches 0; Indels 0; Gaps 0;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 60
DB 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 60

QY 61 AUUGGUUUUAGCAUAGCCUUAUAGGACAGC 90
DB 61 AUUGGUUUUAGCAUAGCCUUAUAGGACAGC 90

RESULT 6
ADJ10053
ID ADJ10053 standard; RNA; 98 BP.
AC ADJ10053;
DT 17-JUN-2004 (first entry)
XX RNA aptamer (RNA9A) SeqID 1.
XX ss; unnatural base; 5'-substituted-2-oxo(1H)-pyridin-3-yl;
KW photoreactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX Unidentified.
XX WO2004067713-A1.
XX 32-JAN-2004.
XX 28-FEB-2003; 2003WO-JP002342.
XX 17-JUL-2002; 2002JP-00208568.
XX (RIKE) RIKEN KK.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX Hirao I, Yokoyama S, Hirao M, Mitsui T;
XX WPI; 2004-122944/12.
XX Nucleosides or nucleotides comprising novel base particularly 5-substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.
XX Example 4; SEQ ID NO 1; 78pp; Japanese.
XX This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)-pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photoreactive group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as base which is applied as template for the integration of any of the nucleosides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification. CC replication and transcription for their preparation. This polynucleotide sequence is an RNA aptamer (RNA9A) given in an exemplification of the invention.
XX Sequence 98 BP; 24 A; 21 C; 24 G; 0 T; 29 U; 0 Other;

Query Match 84.4%; Score 84.4; DB 12; Length 98;
Best Local Similarity 97.0%; Pred. No. 8.9e-20; Mismatches 1; Indels 2; Gaps 1;
Matches 97; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 60
DB 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 58

QY 61 AUUGGUUUUAGCAUAGCCUUAUAGGACAGCAAGCUUCG 100
DB 59 AUUGGUUUUAGCAUAGCCUUAUAGGACAGCAAGCUUCG 98

RESULT 7
AAZ99051
ID AAZ99051 standard; RNA; 80 BP.
XX AAZ99051;
XX 21-JUN-2000 (first entry)
XX RNA aptamer #28 for binding Ras target protein.
XX Ras target protein; malignant tumour; signal transduction regulation;
KW cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX Homo sapiens.
XX WO200009684-A1.
XX 24-FEB-2000.
XX 13-AUG-1999; 99WO-JP004399.
XX 14-AUG-1998; 98JP-00242596.
XX 24-NOV-1998; 98JP-00333284.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Yokoyama S, Hirao I, Sakamoto K;
XX WPI; 2000-224330/19.
XX Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras target protein like Raf-1, useful in drug compositions to treat and diagnose malignant tumors and inflammation.
XX Claim 6; Page 42; 59pp; Japanese.
XX The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99051-299051) based on these sequences are useful in the treatment and diagnosis of malignant tumors and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumors and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal causing proliferation or differentiation of cells
XX Sequence 80 BP; 20 A; 16 C; 19 G; 0 T; 25 U; 0 Other;

Query Match 80.0%; Score 80; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 3e-18; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 60
DB 1 GGGAGUGGAGGAUAUUCGAGGCAUAUGUGGACUCCGUCUCCUUAACAGUUAUA 60

QY 61 AUUGGUUUUAGCAUAGCCU 80
DB 61 AUUGGUUUUAGCAUAGCCU 80

RESULT 8
ADJ10066
ID ADJ10066 standard; DNA; 147 BP.
XX
XX ADJ10066;
XX
XX 17-JUN-2004 (first entry)
XX
XX Synthetic RNA OC aptamer (100-mer).
XX
XX ss; unnatural base; 5'-substituted-2-oxo(1H)-pyridin-3-yl;
KW photoreactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
XX
OS Synthetic.
XX WO2004007713-A1.
XX 22-JAN-2004.
XX
XX 28-FEB-2003; 2003WO-JP002342.
XX
XX 17-JUL-2002; 2002JP-00208568.
XX
XX (RIKE) RIKEN KK.
PA (NISCC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Hirao I, Yokoyama S, Hirao M, Mitsui T;
PI
PI
PI
XX
XX WPI; 2004-122944/12.
XX
XX Nucleosides or nucleotides comprising novel base particularly 5-
PT substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT diseases.
XX
XX
XX Disclosure; Fig 10; 78pp; Japanese.
XX
XX This invention relates to nucleosides or nucleotides that a novel
CC unnatural base, namely the 5'-substituted-2-oxo(1H)-pyridin-3-yl group
CC as a base. Specifically, it refers to nucleic acid molecules integrated
CC with a base substituted at the 5-position with either a photoreactive
CC group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group
CC biotin or a derivative thereof or a fluorescent molecule selected from
CC fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC derivatives thereof. The present invention describes preparing nucleic
CC acids by carrying out transcription, replication or reverse transcription
CC with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC base which is applied as template for the integration of any of the
CC nucleotides into its complementary position. Accordingly, the nucleosides
CC and nucleotides are useful in developing functional nucleic acids
CC including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC therapeutics in treating diseases by gene therapy. The produced nucleic
CC acids are chemically stable with improved ease of amplification,
CC replication and transcription for their preparation. This polynucleotide
CC sequence is a synthetic RNA OC aptamer (100-mer) given in an
XX exemplification of the invention.
XX
SQ Sequence 147 BP; 32 A; 34 C; 37 G; 0 T; 44 U; 0 Other;
Query Match 50.6%; Score 50.6; DB 12; Length 147;
Best Local Similarity 75.2%; Pred. No. 9e-08; 24; Indels 1; Gaps 1
Matches 76; Conservative 0; Mismatches
QY 1 GGGAGUGGAGGAAUUCGAGGAGCAUUGCGACUCCGUCUUCUUAACACAGU-UAAU 59
Db 1 GGGAGUGGAGGAAUUCGAGGAGCAUUGCGACUCCGUCUUCUUAACAGUUAU 60
QY 60 AAUUGGUUUUAGCAUUGCCUUAGCGACGACGAGCUUCUGC 100
Db 61 CACCUGUAAACAGCAUUGCCUUAGCGACGAGCUUCUGC 101


```
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

QY 18 UCGAGGCAUAGUGGACUCCGUCUCCUUAACACAGUUUAUAAUUGGUUUUAGCAUUG 77
Db 317 TTGCGGTGACCTCGACGCGCAGCGTCTTTCAGACCGTGTGATGAACGCGTTGAGTGCATG 258
QY 78 CCUAGCGACGACGCAAGC 94
Db 257 CCCTGGGCACGGGAAGC 241

RESULT 12
ID AAS59583
XX AC AAS59583;
XX DT 13-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein encoding DNA #78.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant; ds.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX PS Claim 1; SEQ ID NO 78; 1069pp; English.
XX CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
XX CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
XX CC associated DNA sequences are used in the treatment, prevention and
XX CC diagnosis of medical conditions caused by P. acnes. The disorders include
XX CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
XX CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
XX CC infections of bone, joints and the central nervous system, however it is
XX CC particularly involved in the inflammatory lesions associated with acne
XX CC vulgaris. A method for detecting the presence or absence of P. acnes in a
XX CC patient comprises contacting a sample with a binding agent that binds to
XX CC the proteins of the invention and determining the amount of bound protein
XX CC in the sample. The polypeptides may be used as antigens in the production
XX CC of antibodies specific for P. acnes proteins. These antibodies can be
XX CC used to downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
XX CC polypeptides shown in AAU57347-AAU57508 and AAU67590-AAU67591. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
```

```
SQ Sequence 15776 BP; 3393 A; 4931 C; 4483 G; 2969 T; 0 U; 0 Other;

Query Match 30.6%; Score 30.6; DB 4; Length 15776;
Best Local Similarity 46.8%; Pred. No. 4.5;
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

QY 18 UCGAGGCAUAGUGGACUCCGUCUCCUUAACACAGUUUAUAAUUGGUUUUAGCAUUG 77
Db 8427 TTGCGGTGACCTCGACGCGCAGCGTCTTTCAGACCGTGTGATGAACGCGTTGAGTGCATG 8486
QY 78 CCUAGCGACGACGCAAGC 94
Db 8487 CCCTGGGCACGGGAAGC 8503

RESULT 13
ID ACF64512
XX AC ACF64512;
XX DT 17-OCT-2003 (first entry)
XX DE Propionibacterium acnes DNA contig sequence #78.
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine; ds.
XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX PA (CORI-) CORIXA CORP.
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX PI Barth B, Vallieve-Douglass J;
XX WPI; 2003-381789/36.
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX PS Claim 1; SEQ ID NO 78; 1481pp; English.
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABW35624-ABW64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
```

CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a P. acnes DNA contig which is specifically claimed
CC in the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 15776 BP; 3393 A; 4931 C; 4483 G; 2969 T; 0 U; 0 Other;

Query Match 30.6%; Score 30.6; DB 8; Length 15776;
Best Local Similarity 46.8%; Pred. No. 4.5; Mismatches 29; Indels 0; Gaps 0;
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;
QY 18 UCGAGCAUAUGUGCAGCUCGCUUCCUCAAACAGUUAUAAUUGGUUAGCAUUG 77
Db 8427 TTGCGGTGACCTCGACGCGCGGGTCTTCAGACCGTTCATGAAACGGTTGAGTCATG 8486

QY 78 CCUAGCGCACGACGC 94
Db 8487 CCTGGCAGCGGAAGC 8503

RESULT 14
ADA29321/c
ID ADA29321 standard; DNA; 366 BP.
AC ADA29321;
XX
XX
DT 20-NOV-2003 (first entry)
XX
DE DNA encoding Acinetobacter baumannii protein #608.

XX ds; Gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW vaccine; plant biocontrol agent.
XX

OS Acinetobacter baumannii.

XX US6562958-B1.

XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;

XX WPI; 2003-576092/54.

XX P-PSDB; ADA33447.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.

XX Example; SEQ ID NO 608; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.

XX Sequence 366 BP; 131 A; 65 C; 63 G; 107 T; 0 U; 0 Other;

Query Match 30.4%; Score 30.4; DB 9; Length 366;
Best Local Similarity 39.8%; Pred. No. 1.7;
Matches 35; Conservative 17; Mismatches 36; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAAUUAUCGAGGCAUAUGUGCAUCUCGCUUCCUCAAACAGUUAUA 60
Db 101 GGGTGTATGATATAGCTTCTATATGTGCACAGTAATGTGCAACAGTAATAG 42
QY 61 AUUGUUUAGCAUAUGCCUUAUGCGACA 88
Db 41 TCTGTTTATGCTATATCTTAGATTCA 14

RESULT 15
AAZ99049
ID AAZ99049 standard; RNA; 98 BP.
XX
AC AAZ99049;
XX
DT 21-JUN-2000 (first entry)
XX
DE RNA aptamer #26 for binding Ras target protein.

XX Ras target protein; malignant tumour; signal transduction regulation;
KW cell proliferation; cell differentiation; aptamer; inflammation; ss.

XX Homo sapiens.

XX WO200009684-A1.

XX 24-FEB-2000.

XX 13-AUG-1999; 99WO-JP004399.

XX 14-AUG-1998; 98JP-00242596.

XX 24-NOV-1998; 98JP-00333284.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yokoyama S, Hirao I, Sakamoto K;

XX WPI; 2000-224330/19.

XX Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT target protein like Raf-1, useful in drug compositions to treat and
PT diagnose malignant tumors and inflammation.

XX Claim 6; Page 41; 59pp; Japanese.

XX The invention relates to novel nucleic acids which bind specifically to
CC the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-299051)
CC based on these sequences are useful in the treatment and diagnosis of
CC malignant tumours and inflammation. The nucleic acids can be used to
CC formulate medicinal compositions that are useful in the treatment of
CC malignant tumours and inflammation and for disease diagnosis by binding
CC specifically to Ras target protein and regulating transmission of signal
CC causing proliferation or differentiation of cells

XX Sequence 98 BP; 24 A; 24 C; 26 G; 0 T; 24 U; 0 Other;

Query Match 30.0%; Score 30; DB 3; Length 98;
Best Local Similarity 63.0%; Pred. No. 1.5;
Matches 63; Conservative 0; Mismatches 35; Indels 2; Gaps 1;

QY 1 GGGAGUGGAGGAAUUAUCGAGGCAUAUGUGCAUCUCGCUUCCUCAAACAGUUAUA 60
Db 1 GGGAGUGGAGGAAUUAUCGAGGCAUAUGUGCAUCUCGCGGAGGUGGUAUUAUUCU 60
QY 61 AUUGUUUAGCAUAUGCCUUAUGCGACAGCAAGCUUCUGC 100
Db 61 CCUACACUU--CUAUGCCUUAUGCGACAGCAAGCUUCUGC 98

Search completed: April 19, 2005, 16:27:31
Job time : 432 secs


```
RESULT 5
US-10-425-114-30801/c
; Sequence 30801, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 28
; SEQ ID NO 30801
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73080F09_FLI
US-10-425-114-30801

Query Match      28.2%; Score 28.2; DB 17; Length 586;
Best Local Similarity 37.1%; Pred. No. 19;
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 8 GAGGAUAUUCGAGGCAUUGUCGACUCCGUCUCCUUAACCAACGUAUUAUUGGUU 67
DB 92 GGGGCAATTTTATGTTGTTATTAGGATTCCTCTGCTCCACCGCATCATAACTCGGTT 33
QY 68 UUAGCAUAUCCUUAUGCGACGACAGCUU 96
DB 32 CCAGCAAGCGCATTTAGGGAAGGAAGATT 4

RESULT 6
US-10-357-930-24310/c
; Sequence 24310, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Endege, wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24310
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24310

Query Match      28.2%; Score 28.2; DB 18; Length 1161;
Best Local Similarity 37.1%; Pred. No. 24;
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 6 UGGAGAAUUCGAGGCAUUGUCGACUCCGUCUCCUUAACCAACGUAUUAUUGGUU 65
DB 878 TGAATAAGCCTTAATGCAATTTCTTCACAGTAATTCAGCCAAATATTTTAAATCAG 819
QY 66 UUUUAGCAUAUCCUUAUGCGACGACAGCAAGC 94
DB 818 GTTAACTTATGCTTAAAGAAATTAAGC 790

RESULT 7
US-10-767-701-13576/c
; Sequence 13576, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13576
; LENGTH: 1612
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS1303_1
US-10-767-701-13576

Query Match      28.2%; Score 28.2; DB 18; Length 1612;
Best Local Similarity 38.2%; Pred. No. 27;
Matches 34; Conservative 17; Mismatches 38; Indels 0; Gaps 0;

QY 8 GAGGAUUCGAGGCAUUGUCGACUCCGUCUCCUUAACCAACGUAUUAUUGGUU 67
DB 1152 GGGGCAATTTTGTGTTGTTATTAGGATTCCTCTGCTCCACTGCGTCACAACTCGGTT 1093
QY 68 UUAGCAUAUCCUUAUGCGACGACAGCAAGCUU 96
DB 1092 CCAGCAAGCGCATTTAGGGAAGGAAGATT 1064

RESULT 8
US-10-425-115-121623/c
; Sequence 121623, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 121623
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_42398C.1
US-10-425-115-121623

Query Match      28.2%; Score 28.2; DB 18; Length 1851;
Best Local Similarity 37.1%; Pred. No. 29;
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 8 GAGGAUUAUCGAGGCAUAUGCAGCUCGCUUCCUUAACACAGUUAUAAUUGGUU 67
DB 1333 GGGGCATTTTATGTTGTTATTAGGATTCCTCCTCCACCGCATCATAACTCGGTT 1274

QY 68 UUAGCAUAUGCCUUAAGCAGACAGCAAGCUU 96
DB 1273 CCAGCAAGCGCATYTAGGGAAGAGAGATT 1245

RESULT 9
US-10-067-514-1/c
; Sequence 1, Application US/10067514
; Publication No. US20030054531A1
; GENERAL INFORMATION:
; APPLICANT: Gretaarsdottir, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reynisdottir, Sigrður Th.
; TITLE OF INVENTION: HUMAN STROKE GENE
; FILE REFERENCE: 2345.2010-003
; CURRENT APPLICATION NUMBER: US/10/067,514
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 09/811/352
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
; TYPE: DNA
; ORGANISM: Human
US-10-067-514-1

Query Match      28.0%; Score 28; DB 14; Length 1691139;
Best Local Similarity 41.7%; Pred. No. 4e+02;
Matches 25; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

QY 12 AAUUAUCGAGGCAUAUGCAGCUCGCUUCCUUAACACAGUUAUAAUUGGUUAG 71
DB 127542 AGTTGATCCATGCTTTTGGAGGAGCTTATCCCTTCAAAGCAGTAATAAAGTGTTTAG 127483

RESULT 10
US-10-419-723-1/c
; Sequence 1, Application US/10419723
; Publication No. US20040014099A1
; GENERAL INFORMATION:
; APPLICANT: Gretaarsdottir, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reynisdottir, Sigrður Th.
; APPLICANT: Thorleifsson, Gudmar
; APPLICANT: Gulcher, Jeffrey
; TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
; TITLE OF INVENTION: METHODS OF TREATMENT
; FILE REFERENCE: 2345.2010-005
; CURRENT APPLICATION NUMBER: US/10/419,723
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 10/255,120
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 10/067,514
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 09/811,352
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-419-723-1

Query Match      28.0%; Score 28; DB 17; Length 1691139;
Best Local Similarity 41.7%; Pred. No. 4e+02;
Matches 25; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

QY 12 AAUUAUCGAGGCAUAUGCAGCUCGCUUCCUUAACACAGUUAUAAUUGGUUAG 71
DB 127542 AGTTGATCCATGCTTTTGGAGGAGCTTATCCCTTCAAAGCAGTAATAAAGTGTTTAG 127483

RESULT 11
US-10-369-493-25760
; Sequence 25760, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25760
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25760

Query Match      27.8%; Score 27.8; DB 17; Length 1093;
Best Local Similarity 38.0%; Pred. No. 33;
Matches 27; Conservative 17; Mismatches 27; Indels 0; Gaps 0;

QY 4 AGUGGAGAAUUAUCGAGGCAUAUGCAGCUCGCUUCCUUAACACAGUUAUAAU 63
DB 213 AGTGGACTCTTTATAGACAATGAAGCAGGATCCGTTTACCTTATAACGAGTTTAAATC 272
QY 64 GGUUUUAGCAU 74
DB 273 GTTTGTTCTAT 283

RESULT 12
US-10-322-281-744/c
; Sequence 744, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 744
; LENGTH: 99291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-744

Query Match      27.8%; Score 27.8; DB 18; Length 99291;
Best Local Similarity 31.0%; Pred. No. 1.7e+02;
Matches 22; Conservative 22; Mismatches 27; Indels 0; Gaps 0;
```

Qy	12	AAUUCACGAGGCAUAUGGCCUCGUCUUCUCAAACCCAGUUUAAAUUGGUUUNG	71
Db	23527	AAATAATGAGACCTGTCAGATACTTTTGGTTTGCACCAATATAAAATGGTTTTAC	23468
Qy	72	CAUAUGCCUUA	82
Db	23467	AAATTGGTTA	23457

RESULT 13

US-10-479-081-104

Sequence 104, Application US/10479081

Publication No. US20050059001A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWARA, AKIRA

TITLE OF INVENTION: NUCLEIC ACIDS ISOLATED FROM NEUROBLASTOMA

FILE REFERENCE: 7388-80893

CURRENT APPLICATION NUMBER: US/10/479,081

CURRENT FILING DATE: 2003-11-26

PRIOR APPLICATION NUMBER: PCT/JP02/05295

PRIOR FILING DATE: 2002-05-30

PRIOR APPLICATION NUMBER: JP 2001-163666

PRIOR FILING DATE: 2001-05-31

PRIOR APPLICATION NUMBER: JP 2001-255260

PRIOR FILING DATE: 2001-08-24

NUMBER OF SEQ ID NOS: 742

SOFTWARE: PatentIn version 3.2

SEQ ID NO 104

LENGTH: 732

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (2)..(2)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (4)..(4)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (15)..(15)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (87)..(87)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (90)..(90)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (104)..(105)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (566)..(566)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (642)..(642)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (650)..(650)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

```

; LOCATION: (706)..(706)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (711)..(711)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (732)..(732)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; OTHER INFORMATION: mbla-03439-f
; US-10-479-081-104

Query Match          27.6%; Score 27.6; DB 19; Length 732;
Best Local Similarity 47.6%; Pred No. 33;
Matches 39; Conservative 9; Mismatches 34; Indels 0; Gaps 0

Qy 10 GGAUAUUGAGGAGCAUAUUGCGUCCGUCUCCAGUUAUAAAUUGUUU 69
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 161 GGAACATCATCCAGACTATCGACACATTGTTCTTCCCGCCACAGTAAGAAAGA 220
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 70 AGCAUAUUGCCUUAAGCGACGCA 91
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 221 TGCAGAAGACTTAGCTGCAGCA 242
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

RESULT 14
US-10-097-111-47
; Sequence 47, Application US/10097111
; Publication No. US20030138771A1
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILLIPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EPI THAT
; TITLE OF INVENTION: ENCODE ANTI-MICROBIAL POLYPEPTIDES
; FILE REFERENCE: 073406-0603
; CURRENT APPLICATION NUMBER: US/10/097,111
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 09/676,412
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/157,218
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 552
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-097-111-47

	Query Match	27.0%	Score 27;	DB 15;	Length 522;
	Best Local Similarity	40.0%	Pred. No. 47;		
	Matches	30;	Mismatches	30;	Gaps 0
Qy	13	AUUAUCGAGGCAUAGUCGACUCGUCUCCUCAAACAGUUAAUUGUUUUUAGC	72		
Db	186	ATTCATTGACAGACTTGACCACCGCTGTTCTTCTTCAAGGGAATGAACCAATCGCTTTAGC	245		
Qy	73	AUUGCCUUAAGCGAC	87		
Db	246	AAATGCAGTTGACAC	260		

RESULT 15
US-10-097-111-10
; Sequence 10, Application US/10097111
; Publication No. US2003013871A1
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILIPPE
; APPLICANT: DUBOW, MICHAEL

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 13:28:09 ; Search time 3095 Seconds
(without alignments)
1229.863 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 999agugaggaaucaucg.....uagocagcagcagcucgc 100

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34.6	34.6	642	2	BE147147	BE147147 PM2-HT022
2	31.4	31.4	749	9	AG557770	AG557770 Mus muscu
3	31.2	31.2	442	8	AZ955882	AZ955882 2M022110
4	30.8	30.8	696	8	B21536	B21536 F21F23-T7 I
5	30.6	30.6	461	6	CB736914	CB736914 AMGNNUC.M
6	30.6	30.6	1074	2	BE371824	BE371824 601217621
7	30.4	30.4	917	9	AG069018	AG069018 Pan trogl
8	30	30	725	7	CK016754	CK016754 AGENCOURT
9	29.8	29.8	528	8	AQ108017	AQ108017 CIT-HSP-2
10	29.6	29.6	578	7	CN553623	CN553623 tae26508
11	29.6	29.6	617	4	BG389990	BG389990 Gm01_01e0
12	29.6	29.6	626	7	CN623811	CN623811 tae59d12
13	29.6	29.6	717	7	CN553289	CN553289 tae26508
14	29.6	29.6	744	7	CN226782	CN226782 RJB004H01
15	29.6	29.6	745	7	CO421969	CO421969 GGEZHT100
16	29.6	29.6	852	5	BU209960	BU209960 603949596
17	29.6	29.6	877	5	BU399017	BU399017 603534483
18	29.2	29.2	254	4	BM401540	BM401540 JH3C07R S
19	29.2	29.2	343	4	BG630802	BG630802 CC-esf1c
20	29.2	29.2	762	7	CF446768	CF446768 EST683113
21	29	29	29.0	715	BZ944697	BZ944697 CH240_121
22	28.8	28.8	339	7	CN464219	CN464219 7866.I AF
C 21	28.8	28.8	532	6	CA249294	CA249294 SCSBFL110
C 23	28.8	28.8	605	9	CE114143	CE114143 tigr-gss-

25	28.8	28.8	739	8	CC115139	CC115139 NDL_81L16
26	28.6	28.6	472	8	AQ955733	AQ955733 LERAF677R
27	28.6	28.6	513	5	BP051024	BP051024 BP051024
C 28	28.6	28.6	559	6	CB094897	CB094897 hz76608.b
C 29	28.6	28.6	561	8	BZ891642	BZ891642 CH240_287
C 30	28.4	28.4	580	8	BZ201271	BZ201271 CH230-351
C 31	28.4	28.4	607	6	CA892404	CA892404 B0172B04-
32	28.4	28.4	629	6	CA890800	CA890800 B0161F04-
33	28.4	28.4	629	6	CA892977	CA892977 B0175H04-
C 34	28.4	28.4	748	9	AG525698	AG525698 Mus muscu
35	28.4	28.4	756	4	BG393101	BG393101 602411369
36	28.4	28.4	818	5	BU708057	BU708057 UI-M-FR0-
C 37	28.4	28.4	832	9	CC491386	CC491386 CH240_325
C 38	28.4	28.4	835	6	CB686108	CB686108 Bn018_03J
C 39	28.2	28.2	489	5	BM594105	BM594105 BM594105
C 40	28.2	28.2	490	5	BM545837	BM545837 BM545837
41	28.2	28.2	496	5	BQ531573	BQ531573 APEX2_5_F
42	28.2	28.2	521	5	BQ531501	BQ531501 APEX3_35
43	28.2	28.2	521	7	CF791691	CF791691 880034 MA
C 44	28.2	28.2	559	4	BG649176	BG649176 EM1_77_E0
C 45	28.2	28.2	598	2	BE363373	BE363373 WSL_62_B0

ALIGNMENTS

RESULT 1
BE147147
LOCUS PM2-HT0224-291099-002-c09 HT0224 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE147147 642 bp mRNA linear EST 21-JUN-2000
ACCESSION BE147147.1 GI:8609871
VERSION BE147147.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 642)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&tl=PM2-HT0224-291099-002-c09&tl=1999-10-29&tl=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 415.

FEATURES

source
Location/Qualifiers
1..642
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0224"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match	34.6%	Score 34.6;	DB 2;	Length 642;
Best Local Similarity	37.0%;	Pred. No. 0.71;		
Matches 30;	Conservative 22;	Mismatches 29;	Indels 0;	Gaps 0;

Qy	12	AATUUAUCGAGGCCAUAUGCAGCCGUCUUCUCAACACAGUUAAUAUUGGUUUUAG	71
	:: :	:: :	: :: :
Db	150	AATAAATTGAGCGCTGTCTCAGATACTTTTGGTTTGCAACAATAATAAATGGTTTTAC	209
	:: :	:: :	: :: :
Qy	72	CAUAUGCCEUAGGCAGCAA	92
	:: :	:: :	:: :
Db	210	AATTTCGGTTTACAACAACAA	230
	:: :	:: :	:: :

RESULT 2	
AG557770/c	
LOCUS	749 bp DNA linear GSS 05-JUN-2004
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-476B14.T7, genomic survey sequence.
ACCESSION	AG557770.1 GI:49318468
VERSION	GSS.
KEYWORDS	
SOURCE	Mus musculus molossinus
ORGANISM	Mus musculus molossinus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Hattori.M., Toyoda.A., Noguchi.H., Kojima.T. and Sakaki.Y.
TITLE	BAC end Sequences of Library MSMg01

REFERENCE	Unpublished
AUTHORS	2 (basses 1 to 749)
TITLE	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y. Direct Submission
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 17-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/ , Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT	Clones are derived from the mouse BAC library MSG01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp

```

PRIMERS
Sequencing : T7
LIBRARY
Vector      : pBACE3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI.
Location/Qualifiers
1. .749
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-476B14.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

```

ORIGIN

Query Match	31.4%;	Score 31.4;	DB 9;	Length 749;
Best Local Similarity	51.0%;	Pred. No.9;		
Matches 25;	Conservative	13;	Mismatches 11;	Indels 0;
				Gaps 0;

Qy	41	UUCCUUCAAACCAAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAAGCGACAG	89
		: : : : : : : : : :	
Db	539	TTCTCTTAAACCAACCGATTATATAAATCTTGCTAGCATATGCAATAGTGCTG	491
RESULT 3			
LOCUS	AZ955882	442 bp	linear
DEFINITION	AZ955882 2M022110F Mouse 10kb plasmid UGC2M library Mus musculus genomic clone UGC2M022110 F, genomic survey sequence.		

ACCESSION	AZ955882	
VERSION	AZ955882.1	GI:13827109
KEYWORDS	GSS.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Eukaryota; Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 442)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0222 row: 1 column: 10
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 442.

```

Location/Qualifiers
1..442
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0222110"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PMW42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMW42 (g14732114[gB/AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

ORIGIN

Query Match	31.2%;	Score 31.2;	DB 8;	Length 442;
Best Local Similarity	32.6%;	Pred. No. 9.8;		
Matches 30;	Conservative 24;	Mismatches 38;	Indels 0;	Gaps 0;

Qy	7	GGAGGAUAUACUAGGAGCAUAGUCGACUCCGCUUACCAACCAAGUUAUAAUUGGU 66
Db	272	GCAGGCTTTTCATCCAGCATCTGATTCCTGAGTCTTCATTGATAAAATCTAACAAATTGAG 331
Qy	67	UUUAGCAUAUAGCCUUVAGCGACGACGAAGCUUCU 98
Db	332	TTTTTGTGTTTGTGTTTTCGAGCAGGTTTCT 363
RESULT 4		
B21536		
LOCUS	B21536	596 bp DNA linear GSS 16-SEP-1997
DEFINITION	F21F23-T7 IGF Arabidopsis thaliana genomic clone F21F23, genomic survey sequence.	
ACCESSION	B21536	
VERSION	B21536	
KEYWORDS	GSS	
SOURCE	B21536.1 GI:2396590	
ORGANISM	Arabidopsis thaliana (thale cress)	
REFERENCE	Arabidopsis thaliana	
AUTHORS	Arabidopsis thaliana	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
	1 (bases 1 to 596)	
	Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.	
TITLE	BAC End Sequences at ATGC	
JOURNAL	Unpublished (1997)	
COMMENT	Other GSSs: F21F23-Sp6	
	Contact: Ecker J.	
	Arabidopsis Thaliana Genome Center	
	University of Pennsylvania	
	Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104	
	Tel: 215-898-9384	
	Fax: 215-898-8780	
	Email: jecker@genome.bio.upenn.edu	
	Seq primer: T7	
	Class: BAC ends	
	High quality sequence start: 129	
	High quality sequence stop: 143.	
FEATURES	Location/Qualifiers	
source	1..596	
	/organism="Arabidopsis thaliana"	
	/mol_type="genomic DNA"	
	/ecotype="Columbia"	
	/db_xref="taxon:3702"	
	/clone="F21F23"	
	/sex="hermaphrodite"	
	/clone_lib="IGF"	
	/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"	
ORIGIN		
Query Match	30.8%; Score:30.8; DB 8; Length 696;	
Best Local Similarity	43.9%; Pred. No.14;	
Matches	29; Conservative 15; Mismatches 22; Indels 0; Gaps 0;	
Qy	9	AGGAUAUACUAGGAGCAUAGUCGACUCCGCUUACCAACCAAGUUAUAAUUGGUU 68
Db	306	AGAAAATCATCGATCAAGTTAGAAAGGTTGTTCTTGACCAACCATATAGATGTGTC 365
Qy	69	UAGCAU 74
Db	366	TACCAT 371
RESULT 5		
CB736914/c		
LOCUS	CB736914	461 bp mRNA linear EST 11-APR-2003
DEFINITION	AGNVCN:MRBE4-00015-C9-A rat brain E15 (10375) Rattus norvegicus cDNA clone mrbe4-00015-c9 5', mRNA sequence.	
ACCESSION	CB736914	
VERSION	CB736914.1 GI:29804127	

/clone="IMAGE:3586661"
 /tissue type="spontaneous tumor, metastatic to mammary.
 Stem cell origin."
 /lab_host="PH10B"
 /clone_lib="NCI_CGAP Lu29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 30.6%; Score 30.6; DB 2; Length 1074;
 Best Local Similarity 44.2%; Pred. No. 18;
 Matches 34; Conservative 14; Mismatches 29; Indels 0; Gaps 0;
 QY 19 CGAGCAUAGUGCAGCUCGCUUUAUAAACCAAGUUAUUAUUGUUUAGCAUAGCC 78
 Db 280 CCAGGATTCAGCAGACCCCTGCTCCCTGAACCCCACTTATTAGTCTGTGAGCATATAC 339
 QY 79 CUUAGCGCAGCAAGCU 95
 Db 340 CATGGGACTTCATGCT 356

RESULT 7
 AG069018/c
 LOCUS 917 bp DNA linear GSS 03-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-059H17.R, genomic survey sequence.
 ACCESSION AG069018
 VERSION AG069018.1 GI:16620820
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of Library PTB
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 917)
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimpsesgsc.riken.go.jp, URL: http://hgp.sec.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI

FEATURES

source
 1..917
 Location/Qualifiers
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-059H17.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 30.4%; Score 30.4; DB 9; Length 917;
 Best Local Similarity 39.1%; Pred. No. 20;
 Matches 25; Conservative 18; Mismatches 21; Indels 0; Gaps 0;

QY 22 GGCAUAUUGCAGCUCGCUUUAUAAACCAAGUUAUUAUUGUUUAGCAUAGCCUU 81
 Db 366 GGTACTGTGTCGCCCTCTCTCTAGCGTGGGGTTATAATGATTAAATATACCTT 307
 QY 82 ACGC 85
 Db 306 CCCG 303

RESULT 8

CK016754/c

LOCUS

DEFINITION AGENCOURT 16544280 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7043617
 5', mRNA sequence.

ACCESSION

CK016754

VERSION

CK016754.1 GI:38541867

KEYWORDS

EST.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 725)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Len Zon, Harvard

cDNA Library Preparation: Open Biosystems

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14803 row: i column: 23

High quality sequence stop: 594.

location/Qualifiers

1..725

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="IMAGE:7043617"

/tissue_type="whole body"

/lab_host="DH10B"

/clone_lib="NIH_ZGC_10"

/note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;

Bulk tissue was collected from a whole adult individual

from the Tuebingen strain. 1st strand cDNA was primed with

a Not I - oligo(dT) primer, double-stranded cDNA was

cloned into the Not I and EcoRV sites of pExpress-1.

Library was size-selected for >1 kb fragments. A

normalized version of this library is also available

(NIH_ZGC_7). Library was constructed by Open Biosystems

(Huntsville, AL)."

ORIGIN

Query Match 30.0%; Score 30; DB 7; Length 725;
 Best Local Similarity 44.3%; Pred. No. 27;
 Matches 31; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 2 GGAGUGGAGGAUUAUUGCAGGCAUUGUUGACUCCGUUUAUUAUUAUUA 61

Db 663 GGAGTGTAAAGATTTCAGCAAGACATGTGGGCATCCGTAGTCTTTGAAGCATATATA 604

QY 62 UUGGUUUUAG 71

Db 603 TAAGTAGTAG 594

QY 81 UAGCGACAGCAGCUU 96
 Db 526 CAACGAGATCAAGTTT 541

RESULT 14
 CN226782
 LOCUS CN226782 744 bp mRNA linear EST 09-APR-2004
 DEFINITION RJB004H01.ab1 Rjtestis Gallus gallus cdna 5', mRNA sequence.
 ACCESSION CN226782
 VERSION CN226782.1 GI:46330526
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 744)
 AUTHORS Savolainen, P., Fitzsimmons, C.J., Arvestad, L., Andersson, L. and Lundberg, J.
 TITLE EST analysis of brain and testis cDNA libraries from White Leghorn and Red Jungle Fowl
 JOURNAL Unpublished (2004)
 COMMENT Contact: Peter Savolainen
 Department of Biotechnology
 Royal Institute of Technology, KTH
 SE-106 91 Stockholm, SWEDEN
 Tel: +46 (0)8 5537 8481
 Fax: +46 (0)8 5537 8335
 Email: Peter.Savolainen@biotech.kth.se
 Seq primer: M13 reverse primer.
 Location/Qualifiers
 1..744
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Red junglefowl"
 /db_xref="taxon:9031"
 /sex="male"
 /lab_host="ElectronMAX DH10B (Invitrogen)"
 /clone_lib="Rjtestis"
 /note="Organ: testis; Vector: pSPORT-1; Site 1: Hind III; Site 2: EcoRI; The cDNA libraries were created with the Superscript Plasmid System (Invitrogen)."

FEATURES
 source
 1..744
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 /tissue_type="pituitary and hypothalamus"
 /dev_stage="21 days old"
 /lab_host="DH10B"
 /clone_lib="HTI"
 /note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed with the SuperScript Plasmid System with Gateway Technology kit (Invitrogen), following manufacture's protocols. Plasmid DNA was purified using a modified alkaline lysis method. Sequencing reactions were conducted using the kit Big Dye Terminator Cycle Sequencing Ready Reaction (Applied Biosystems) according to the manufacturer's recommendations. Clones were sequenced by the 5' end with T7 primer. Sequencing reactions were analyzed on ABI Prism 3100 Genetic Analyzer (Applied Biosystems). The quality and clustering of the ESTs were analyzed using the softwares Phred/Gap3. Only EST sequences with Phred quality greater than 20 and at least 150 bp were considered for clustering."

ORIGIN
 Query Match 29.6%; Score 29.6; DB 7; Length 744;
 Best Local Similarity 55.8%; Pred. No. 37;
 Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAUAUCGAGGCAUAUCGACUCCGUCUUCUUAACCC 52
 ||||| :
 Db 254 GGGAGTTTAAAGGCTTCATCGAGCAATGGGAGCTCGTCTCCATCGCACC 305
 ||||| :
 Search completed: April 19, 2005, 16:20:16
 Job time : 3104 secs

QY 81 UAGCGACAGCAGCUU 96
 Db 526 CAACGAGATCAAGTTT 541

RESULT 15
 CN226782
 LOCUS CN226782 745 bp mRNA linear EST 02-JUL-2004
 DEFINITION GGEZHT1005A10.g HT1 Gallus gallus cdna clone GGEZHT1005A10, mRNA sequence.
 ACCESSION CO421969
 VERSION CO421969.1 GI:49638217
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 745)
 AUTHORS Silva, C.S., Jorge, E.C., Patricio, M., Ledur, M.C. and Coutinho, L.L.
 TITLE Discovery of new genes expressed in the chicken pituitary and hypothalamus
 JOURNAL Unpublished (2004)
 COMMENT Contact: Clarissa S. Silva

Laboratory of Animal Biotechnology, Dep. of Animal Production
 ESALQ - University of Sao Paulo
 Av. Padua Dias, 11, Piracicaba, SP, 13419-900, Brazil
 Tel: 55 19 3429 4434
 Fax: 55 19 3429 4285
 Email: cssilva@esalq.usp.br and llicoutin@esalq.usp.br
 PCR Primers
 BACKWARD: T7.

FEATURES
 source
 1..745
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 /clone_lib="GGEZHT1005A10"
 /tissue_type="pituitary and hypothalamus"
 /dev_stage="21 days old"
 /lab_host="DH10B"
 /clone_lib="HTI"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed with the SuperScript Plasmid System with Gateway Technology kit (Invitrogen), following manufacture's protocols. Plasmid DNA was purified using a modified alkaline lysis method. Sequencing reactions were conducted using the kit Big Dye Terminator Cycle Sequencing Ready Reaction (Applied Biosystems) according to the manufacturer's recommendations. Clones were sequenced by the 5' end with T7 primer. Sequencing reactions were analyzed on ABI Prism 3100 Genetic Analyzer (Applied Biosystems). The quality and clustering of the ESTs were analyzed using the softwares Phred/Gap3. Only EST sequences with Phred quality greater than 20 and at least 150 bp were considered for clustering."

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 745;
 Best Local Similarity 55.8%; Pred. No. 37;
 Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAUAUCGAGGCAUAUCGACUCCGUCUUCUUAACCC 52
 ||||| :
 Db 254 GGGAGTTTAAAGGCTTCATCGAGCAATGGGAGCTCGTCTCCATCGCACC 305
 ||||| :
 Search completed: April 19, 2005, 16:20:16
 Job time : 3104 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 13:28:09 ; Search time 130 Seconds
(without alignments)
1258.674 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggagugaggaaucaugc.....uagcgacgaagcuucgc 100

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	30.4	30.4	366	US-09-328-352-608	Sequence 608, Appl
2	29.8	29.8	4285	US-09-410-464-1	Sequence 1, Appli
3	27.4	27.4	405	US-09-270-767-6441	Sequence 6441, Ap
4	27.4	27.4	405	US-09-270-767-21723	Sequence 21723, A
C 5	27.4	27.4	3791	US-08-675-773B-3	Sequence 3, Appli
C 6	27.4	27.4	260286	US-09-949-016-17037	Sequence 17037, A
C 7	27.4	27.4	260293	US-09-949-016-12106	Sequence 12106, A
C 8	27.0	27.0	208	US-09-513-999C-21500	Sequence 21500, A
C 9	26.8	26.8	415	US-09-513-999C-26061	Sequence 26061, A
10	26	26	6944	US-09-949-016-3637	Sequence 3637, Ap
C 11	26	26	106380	US-09-949-016-17553	Sequence 17553, A
12	26	26	150597	US-09-949-016-15379	Sequence 15379, A
13	25.8	25.8	798	US-08-956-171E-162	Sequence 162, App
14	25.8	25.8	798	US-08-781-986A-162	Sequence 162, App
15	25.8	25.8	77851	US-09-949-016-12508	Sequence 12508, A
16	25.8	25.8	77867	US-09-949-016-13211	Sequence 13211, A
17	25.8	25.8	77867	US-09-949-016-13212	Sequence 13212, A
18	25.8	25.8	77940	US-09-949-016-12509	Sequence 12509, A
19	25.6	25.6	372	US-08-956-171E-798	Sequence 798, App
20	25.6	25.6	372	US-08-781-986A-798	Sequence 798, App
21	25.6	25.6	27702	US-09-949-016-11795	Sequence 11795, A
22	25.4	25.4	3711	US-09-902-540-7049	Sequence 7049, Ap
C 23	25.4	25.4	13624	US-09-902-540-1053	Sequence 1053, Ap
C 24	25.4	25.4	50229	US-09-949-016-15499	Sequence 15499, A
25	25.4	25.4	276687	US-09-949-016-13840	Sequence 13840, A
26	25.2	25.2	355	US-09-513-999C-8830	Sequence 8830, Ap
27	25.2	25.2	601	US-09-949-016-190077	Sequence 190077,

C	28	25.2	25.2	18572	4	US-09-949-016-17183	Sequence 17183, A
	29	25.2	25.2	57178	4	US-09-949-016-12838	Sequence 12838, A
	30	25.2	25.2	57178	4	US-09-949-016-17458	Sequence 17458, A
	31	25.2	25.2	786431	4	US-09-751-389-3	Sequence 3, Appl
	32	25	25.0	246	4	US-09-513-999C-35958	Sequence 35958, A
	33	25	25.0	960	4	US-09-328-352-1916	Sequence 1916, Ap
	34	25	25.0	47781	4	US-09-949-016-16492	Sequence 16492, A
	35	25	25.0	47781	4	US-09-949-016-16493	Sequence 16493, A
	36	25	25.0	47781	4	US-09-949-016-16494	Sequence 16494, A
	37	25	25.0	1830121	4	US-09-557-884-1	Sequence 1, Appl
	38	25	25.0	1830121	4	US-09-643-990A-1	Sequence 12, Appl
	39	24.8	24.8	577	4	US-09-706-228-12	Sequence 4695, Ap
	40	24.8	24.8	2775	4	US-09-248-796A-4695	Sequence 42, Appl
	41	24.8	24.8	3786	3	US-08-975-762-42	Sequence 42, Appl
	42	24.8	24.8	3786	3	US-09-295-028-42	Sequence 42, Appl
	43	24.8	24.8	3786	3	US-09-106-582-42	Sequence 42, Appl
	44	24.8	24.8	3786	4	US-09-159-469-42	Sequence 42, Appl
	45	24.8	24.8	3786	4	US-09-693-542-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-328-352-608/c
; Sequence 608, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 608
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-608

Query Match 30.4%; Score 30.4; DB 4; Length 366;
Best Local Similarity 39.8%; Pred. No. 0.2;
Matches 35; Conservative 17; Mismatches 36; Indels 0; Gaps 0;

Qy	1	GGGAGGAGGAGAAUUAUCGAGGCAUAUGGCAUCCGUCUCCUUAACCAAGUAUAA 60
Db	101	GGGTGTTTATGATATAGCTTCTATATGTCGACCAAGTAATGTGCAACAAAGTAATAG 42
Qy	61	AUUGGUUUUAGCAUAGCCUUAAGCGACA 88
Db	41	TCGTGTTTATGCGTATATCTTAGATTCA 14

RESULT 2
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4285

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2005, 13:28:09 ; Search time 1693 Seconds
(without alignments)
2862.092 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaaucaugc.....uagcgacagcaagcuuugc 100

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ste.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32.2	32.2	199415	10 AC084053	AC084053 Mus muscu
C 2	31.2	31.2	201197	10 AC125279	AC125279 Mus muscu
C 3	30.8	30.8	187606	2 AC110847	AC110847 Rattus no
C 4	30.8	30.8	221981	2 AC103165	AC103165 Rattus no
C 5	30.8	30.8	231814	2 AC114165	AC114165 Rattus no
C 6	30.8	30.8	233823	2 AC110697	AC110697 Rattus no
C 7	30.8	30.8	238172	2 AC099144	AC099144 Rattus no
C 8	30.8	30.8	250740	2 AC129051	AC129051 Rattus no
C 9	30.6	30.6	748	6 CQ451733	CQ451733 Sequence
C 10	30.6	30.6	15776	6 CQ363795	CQ363795 Sequence
C 11	30.6	30.6	110000	1 AE017283_15	Continuation (16 o
C 12	30.4	30.4	217	4 AF522912	AF522912 Tachyglos
C 13	30.4	30.4	217	4 AF522914	AF522914 Tachyglos
C 14	30.4	30.4	217	4 AF522916	AF522916 Tachyglos
C 15	30.4	30.4	366	6 AR318058	AR318058 Sequence
C 16	30.4	30.4	3185	4 AV194920	AV194920 Tachyglos
C 17	30.4	30.4	46077	2 AC137289	AC137289 Rattus no
C 18	30.4	30.4	58143	2 AC137177	AC137177 Rattus no
C 19	30.4	30.4	207260	2 AC134220	AC134220 Rattus no

20	30.4	30.4	213033	2	AC131482	AC131482 Rattus no
C 21	30.4	30.4	218157	2	AC139588	AC139588 Rattus no
C 22	30.4	30.4	272400	2	AC110146	AC110146 Rattus no
C 23	30.4	30.4	283299	2	AC112576	AC112576 Rattus no
C 24	30.2	30.2	191674	4	CR536601	CR536601 Platypus
C 25	30.2	30.2	229522	2	AC109911	AC109911 Rattus no
C 26	30.2	30.2	235937	2	AC131223	AC131223 Rattus no
C 27	30.2	30.2	240755	2	AC095368	AC095368 Rattus no
C 28	30	30	207191	2	AC133102	AC133102 Mus muscu
C 29	29.8	29.8	4285	6	AR372457	AR372457 Sequence
C 30	29.8	29.8	4285	8	AF057708	AF057708 Populus b
C 31	29.8	29.8	184147	9	AC103805	AC103805 Homo sapi
C 32	29.8	29.8	202287	2	AC145887	AC145887 Pan trogl
C 33	29.8	29.8	215994	9	AC067941	AC067941 Homo sapi
C 34	29.8	29.8	223143	2	CR450730	CR450730 Danio rer
C 35	29.8	29.8	227025	2	AC148836	AC148836 Pan trogl
C 36	29.6	29.6	23393	2	AC018216	AC018216 Drosophil
C 37	29.6	29.6	50481	2	AC117541	AC117541 Mus muscu
C 38	29.6	29.6	79826	3	AC002443	AC002443 Drosophil
C 39	29.6	29.6	110000	2	AC108332_3	Continuation (4 of
C 40	29.6	29.6	110000	2	AC12373_04	Continuation (5 of
C 41	29.6	29.6	156806	3	AC093047	AC093047 Drosophil
C 42	29.6	29.6	182816	5	AC145913	AC145913 Gallus ga
C 43	29.6	29.6	192132	3	AC093440	AC093440 Drosophil
C 44	29.6	29.6	234627	2	AC106118	AC106118 Rattus no
C 45	29.6	29.6	237005	2	AC137343	AC137343 Rattus no

ALIGNMENTS

RESULT 1	AC084053	199415 bp	DNA	linear	ROD 05-SEP-2002
LOCUS	AC084053	Mus Musculus Strain C57BL6/J chromosome 3 BAC, RP23-158F1, complete sequence.			
DEFINITION	AC084053	HTG.			
ACCESSION	AC084053	GI:27275947			
VERSION	AC084053.6				
KEYWORDS	HTG.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.				
TITLE	High Throughput Mouse Sequencing				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 199415)				
AUTHORS	Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-OCT-2000) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA				
REFERENCE	3 (bases 1 to 199415)				
AUTHORS	Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JUL-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA				
REFERENCE	4 (bases 1 to 199415)				
AUTHORS	Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA				


```

repeat_region 23882..24118
                /rpt_family="Lx4"
repeat_region complement(24228..24641)
                /rpt_family="Lx6"
repeat_region complement(24821..25092)
                /rpt_family="Lx3"
repeat_region 25183..25208
                /rpt_family="Lx3"
repeat_region 25210..25277
                /rpt_family="Lx3"
repeat_region complement(25288..25661)
                /rpt_family="Lx3"
repeat_region complement(26114..32348)
                /rpt_family="Lx3"

Query Match      32.2% Score 32.2; DB 10; Length 199415;
Best Local Similarity 49.1%; Pred. No. 8.6;
Matches 26; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 37 CGUCUCCUUAACACAGUUAUUAUGGUUUUAGCAUUGCCUAGCGACAG 89
Db 63221 CTTATTCCTTAAACACAGTATAAATCTTGCTAGCATATGCAATAGTGTCTG 63273

```

```

RESULT 2
AC125279/c
LOCUS          201197 bp DNA linear ROD 05-NOV-2003
DEFINITION    Mus musculus BAC clone RP23-433F5 from 8, complete sequence.
ACCESSION     AC125279
VERSION       AC125279.1 GI:21536173
KEYWORDS      HTG.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
               1 (bases 1 to 201197)
               Trani, L. and Cotton, M.
               The sequence of Mus musculus BAC clone RP23-433F5
               Unpublished (2001)
               2 (bases 1 to 201197)
               Wilson, R.
               Sequencing of Mus musculus
               Unpublished (2001)
               3 (bases 1 to 201197)
               McPherson, J.D. and Waterston, R.H.
               Direct Submission
               Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
               Parkway, St. Louis, MO 63108, USA
               4 (bases 1 to 201197)
               Wilson, R.
               Direct Submission
               Submitted (05-NOV-2003) Department of Genetics, Washington
               University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
               ----- Genome Center
               Center: Washington University Genome Sequencing Center
               Center code: WUGSC
               Web site: http://genome.wustl.edu
               Contact: submissions@watson.wustl.edu
               ----- Summary Statistics
               -----
               Center project name: M_BA0433F05

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tatenio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES	source	Location/Qualifiers
repeat_region	1..201197	/organism="Mus musculus"
repeat_region	171..557	/mol_type="genomic DNA"
repeat_region	558..1352	/db_xref="taxon:10090"
repeat_region	1353..1499	/chromosome="8"
repeat_region	1500..1886	/map="8"
repeat_region	1887..2101	/clone="RP23-433F5"
repeat_region	2103..2485	/clone_lib="RPCI-23"
repeat_region	2792..2918	/rpt_family="MaLR"
repeat_region	2916..3397	/rpt_family="MaLR"
repeat_region	3398..3556	/rpt_family="Alu"
repeat_region	3557..6980	/rpt_family="MaLR"
repeat_region	6981..7220	/rpt_family="ERV"
repeat_region	7221..8262	/rpt_family="MaLR"
repeat_region	8263..9125	/rpt_family="L1"
repeat_region	9126..10473	/rpt_family="L1"
repeat_region	10474..10619	/rpt_family="L1"
repeat_region	10620..10940	/rpt_family="MER1_type"
repeat_region	10941..11051	/rpt_family="ERV"
repeat_region	11052..12140	/rpt_family="RMR17C"
repeat_region	12141..12714	/rpt_family="L1"
repeat_region	12715..12907	/rpt_family="L1"
repeat_region	12908..13742	/rpt_family="B2"
repeat_region	13743..14342	/rpt_family="L1"

```

repeat_region 14760..15436
/rpt_family="Li"
repeat_region 16028..16131
/rpt_family="MaLR"
repeat_region 17836..18283
/rpt_family="Li"
repeat_region 18352..18653
/rpt_family="ERVK"
repeat_region 18655..19078
/rpt_family="ERVK"
repeat_region 19375..19942
/rpt_family="Li"
repeat_region 21030..21117
/rpt_family="MaLR"
repeat_region 22308..22444
/rpt_family="MIR"
repeat_region 22775..22988
/rpt_family="MER1_type"
repeat_region 22991..23088
/rpt_family="Li"
repeat_region 23274..23681
/rpt_family="MaLR"
repeat_region 24784..24971
/rpt_family="B2"
repeat_region 25314..25486
/rpt_family="MaLR"
repeat_region 26145..26223
/rpt_family="RAL_RN"
repeat_region 26473..26619
/rpt_family="Alu"
repeat_region 26817..27001
/rpt_family="B2"
tRNA complement(26924..26996)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=39.59 / Sec struct
Sc=-11.21)"
repeat_region 27016..27079
/rpt_family="Alu"
repeat_region 28465..28604
/rpt_family="Alu"
repeat_region 28818..29133
/rpt_family="MaLR"
repeat_region 29251..29389
/rpt_family="MIR"
repeat_region 29545..29687
/rpt_family="MaLR"
repeat_region 30593..30737
/rpt_family="L2"
repeat_region 32073..32437
/rpt_family="MaLR"
repeat_region 33137..33276
/rpt_family="Alu"
repeat_region 33402..33547
/rpt_family="Alu"
repeat_region 35207..35290
/rpt_family="MIR"
repeat_region 35653..35830
/rpt_family="MaLR"
repeat_region 35831..36039
/rpt_family="B2"
repeat_region 36040..36232
/rpt_family="MaLR"
repeat_region 36324..36549
/rpt_family="B4"
repeat_region 36709..36788
/rpt_family="B4"
repeat_region 36795..36891
/rpt_family="ERVL"
repeat_region 36862..37099
/rpt_family="MaLR"
repeat_region 39693..39806
/rpt_family="MER1_type"
repeat_region 40903..41102

```

```

/rpt_family="MER2_type"
41131..41296
/rpt_family="MER2_type"
41519..41626
/rpt_family="L2"
42585..42730
/rpt_family="Alu"

Query Match 31.2%; Score 31.2; DB 10; Length 201197;
Best Local Similarity 32.6%; Pred. No. 19;
Matches 30; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

QY 7 GGAGGAUUUCACGAGGCAUUGCAGUCGUCUUCUACCAACCAAGUUAUUGGU 66
Db 2984 GCAGGCTTTTCATCCAGCATCTGATTCCTGAGTCTTCATGATAAATCTAACAATTGAG 2925
QY 67 UUUAGCAUUGCCUAGGACGAGCAAGCUUCU 98
Db 2924 TTTTGTGTTTGTGTTTTCGAGACCAAGGTTTCT 2893

RESULT 3
AC110847
LOCUS
DEFINITION
AC110847 187606 bp DNA linear HTG 15-NOV-2002.
Rattus norvegicus clone CH230-42M10, *** SEQUENCING IN PROGRESS
**, 13 unordered pieces.
AC110847 GI:25006792
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 187606)
Murny,D.Marie., Metzker,M.Iee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Liu,J.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackeleme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

```


Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 233823)

Muzny,D,Marie, Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Joliviet,A., Karpachy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., Liu,W., Longacre,S., Lopez,J., Lorensheva,L., Louised,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K., Nwaokelumen,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sleson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steidle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 233823)

Worley,K.C.

Direct Submission

Submitted (15-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 233823)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 23, 2002 this sequence version replaced gi:23267238.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRXR
Center clone name: CH230-43H7

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 212738 bases at least Q40
Consensus quality: 215401 bases at least Q30
Consensus quality: 217067 bases at least Q20
Estimated insert size: 219523; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 233823: contig of 233823 bp in length.

FEATURES

source
1. 233823
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-43H7"
1. 1276
/note="wgs end extension
clone_end:T7"
1327..2821
/note="wgs end extension
clone_end:T7"
7344..8215
/note="clone boundary
clone_end:T7"
site:EcoRI
end_sequence:BH264888"
85851..88514
/note="wgs contig"
complement(232161..232332)
/note="clone boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH264890"

misc_feature
1327..2821
/note="wgs end extension
clone_end:T7"
7344..8215
/note="clone boundary
clone_end:T7"
site:EcoRI
end_sequence:BH264888"
85851..88514
/note="wgs contig"
complement(232161..232332)
/note="clone boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH264890"

ORIGIN

Query Match 30.8%; Score 30.8; DB 2; Length 233823;
Best Local Similarity 34.7%; Pred. No. 26;
Matches 34; Conservative 22; Mismatches 42; Indels 0; Gaps 0;
QY 2 GGAGUGGAGGAUAUUGCAGGCAUAUGCAGCUCCUUCUUAACACGUAUAAA 61
DB 230212 GTAGAGCAGATATCATTTAGGAGTTATTATCACTGTTTCTTAAAGTCATTATTAT 230271
QY 62 UUGGUUUGCAUUGCCUUAUGCGCAGCAGCAAGCUUCUG 99
DB 230272 TTAGTTTTCCTAGTCTCTTGGGCCCATCAAGTTTCAG 230309

Estimate: www.collegeboard.com/sat

ORIGIN

Query Match	30.8%;	Score 30.8;	DB 2;	Length 250740;
Best Local Similarity	43.1%;	Pred. No. 27;		
Matches	25;	Conservative	16; Mismatches	17; Indels 0; Gaps 0;

QY	23	GCAUAUGCAGUCGCGUCCUUCAAACAGAUUAAUUGGUUUUAGCAUAUGCUC	80
Dd	4082	GCATATTTCATCTTCTTGCCCTTAACAACAAATTGAACGTGGTCTTACAAAATCTCT	4025

```

RESULT 9
CQ451733/c
LOCUS          748 bp      DNA
DEFINITION     Sequence 17493 from Patent WO0192523.
ACCESSION      CQ451733
VERSION         CQ451733.1 GI:41420097
KEYWORDS
SOURCE          Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS        Shimkets,R.A. and Leach,M.D.
TITLE          Human polynucleotides and polypeptides encoded thereby
JOURNAL        Patent: WO 0192523-A 17493 06-DEC-2001;
                Curagen Corporation (US)
FEATURES
source         Location/Qualifiers
                1..748
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

```

Query Match	30.6%	Score 30.6;	DB 6;	Length 748;
Best Local Similarity	46.8%;	Pred.No.20;		
Matches	36;	Conservative	12;	Mismatches 29; Indels 0; Gaps 0;
Qy	18	UCGAGGCAUAUGUCACUCCGCUCUCCUCUCAAAACACAGUIAUAUAUUGUUUAGCAUAUG	77	
Dd	317	TTGCGTGGCAGCTCGACGCCAGCGTTCTTCAGACC GTTGATGAACCGTTGTAGTGCATG	258	
Qy	78	CCUAGCGCACGCAAGC	94	
Dd	257	CCCTGGGCACGGGAAGC	241	

RESULT 10
 CQ363795
 LOCUS linear DNA 15776 bp
 DEFINITION Sequence 78 from Patent WO0181581.
 ACCESSION CQ363795
 VERSION CQ363795.1 GI:41300489
 KEYWORDS
 SOURCE
 ORGANISM
 PROPRIONIBACTERIUM ACNES
 PROPRIONIBACTERIUM ACNES
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Propionibacterineae; Propionibacteriaceae; Propionibacterium.
 1
 AUTHORS Skelly, Y.A., Persing, D.H., Mitcham, J.L., Wang, S.S., Bhatia, A.,
 L'Maisonmeuve, J.F., Zhang, Y., Jen, S. and Carter, D.
 TITLE Compositions and methods for the therapy and diagnosis of acne
 vulgaris
 JOURNAL Patent: WO 0181581-A 78 01-NOV-2001;
 CORIXA CORPORATION (US)
 FEATURES
 source 1. .15776
 /organism="Propionibacterium acnes"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1747"
 ORIGIN

	Query Match	30.8%	Score 30.6	DB 6	Length 15776
	Best Local Similarity	46.8%	Pred. No. 25		
	Matches 36	Conservative 12	Mismatches 29	Indels 0	Gaps 0
Qy	18	UCGAGGCAUAGUCGACUCGCGUCUCUUCUCAAACCGAUUAAAAUUUGUUUUUAGCAUAUG	77		
Db	8427	TTGGGTCGACCTCGACGCGCGGTTCTTTCAGACCGTTGTGTAACGGTTGTAGGTCATG	8486		
Qy	78	CCUUAGCGCAGCAAGC	94		
Db	8487	CCCTGGGCACGGGAAGC	8503		

RESULT 11
AE017283.15
WPCOMMENT
Sequence split into 26 fragments LOCUS AE017283 Accession AE017283

Fragment Name	Begin	End
AE017283_00	1	110000
AE017283_01	100001	210000
AE017283_02	200001	310000
AE017283_03	300001	410000
AE017283_04	400001	510000
AE017283_05	500001	610000
AE017283_06	600001	710000
AE017283_07	700001	810000
AE017283_08	800001	910000
AE017283_09	900001	1010000
AE017283_10	1000001	1110000
AE017283_11	1100001	1210000
AE017283_12	1200001	1310000
AE017283_13	1300001	1410000
AE017283_14	1400001	1510000
AE017283_15	1500001	1610000
AE017283_16	1600001	1710000
AE017283_17	1700001	1810000
AE017283_18	1800001	1910000
AE017283_19	1900001	2010000
AE017283_20	2000001	2110000
AE017283_21	2100001	2210000
AE017283_22	2200001	2310000
AE017283_23	2300001	2410000
AE017283_24	2400001	2510000
AE017283_25	2500001	2560265

	Query Match	30.8%;	Score 30.6;	DB 1;	Length 110000;
	Best Local Similarity	46.8%;	Pred. No. 29;		
	Matches	36;	Conservative 12;	Mismatches 29;	Indels 0; Gaps 0;
Qy	18	UCGAGGCAUUGUCGACUCGCUUCUUCACCAACGAGUUAUAAAUGUUUUAAGCAUUG	77		
Db	34666	TTCCGGTCGACCTCGACGCGCAGCGTTCTTCAGACCGTTGTATGAAACGGTTGTAGTCATG	34722		
Qy	78	CCUUAGGGACAGCAAGC	94		
Db	34726	CCCTGGGCGACGGGAAGC	34742		

RESULT 12					
AF522912/c					
LOCUS	AF522912	217 bp	DNA	linear	MAM 22-MAY-2003
DEFINITION	Tachyglossus aculeatus clone Tag7 type I interferon gene, partial cds.				
ACCESSION	AF522912				
VERSION	AF522912.1	GI:27451581			
KEYWORDS	.				
SOURCE	Tachyglossus aculeatus (Australian echidna)				
ORGANISM	Tachyglossus aculeatus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Monotremata; Tachyglorissidae; Tachyglossus.				
REFERENCE	(base 1 to 217)				

THIS PAGE BLANK (USPTO)